

# An extremal problem in coloring of hypergraphs

Tapas Kumar Mishra   Sudebkumar Prasant Pal  
Dept. of Computer Science and Engineering  
IIT Kharagpur 721302, India

Friday 10<sup>th</sup> July, 2015

## Abstract

Let  $G(V, E)$  be a  $k$ -uniform hypergraph. A hyperedge  $e \in E$  is said to be properly  $(r, p)$  colored by an  $r$ -coloring of vertices in  $V$  if  $e$  contains vertices of at least  $p$  distinct colors in the  $r$ -coloring. An  $r$ -coloring of vertices in  $V$  is called a *strong  $(r, p)$  coloring* if every hyperedge  $e \in E$  is properly  $(r, p)$  colored by the  $r$ -coloring. We study the maximum number of hyperedges that can be properly  $(r, p)$  colored by a single  $r$ -coloring and the structures that maximizes number of properly  $(r, p)$  colored hyperedges.

**Keywords:** Hypergraph, coloring, Strong coloring, extremal problem

## 1 Introduction

Let  $G(V, E)$  be a  $n$ -vertex  $k$ -uniform hypergraph. A hyperedge  $e \in E$  is *properly  $(r, p)$  colored* by an  $r$ -coloring of vertices if  $e$  consists of at least  $p$  distinctly colored vertices. A *strong  $(r, p)$  coloring* of  $G$  is an  $r$ -coloring of the vertices of  $V$ , such that  $\forall e \in E$ ,  $e$  consists of at least  $p$  distinctly colored vertices. We note that for a fixed  $r$  and  $p$ ,  $G$  may not have any strong  $(r, p)$  coloring. Moreover, its not too hard to see that the decision problem is also *NP*-complete, since (i) the decision problem of bicolorability of hypergraphs is *NP*-complete ([7]), and (ii) proper  $(2, 2)$  coloring of  $G$  is equivalent to proper bicoloring of  $G$ . Given  $n$ ,  $k$ ,  $r$ , and  $p$ , we study the maximum number of hyperedges  $M(n, k, r, p)$  of any  $n$ -vertex  $k$ -uniform hypergraph  $G(V, E)$  that can be properly  $(r, p)$  colored by a single  $r$ -coloring.

This problem has an equivalent counterpart in graphs. A proper coloring of a edge in graphs denotes the vertices of the edge getting different colors. A graph is properly colored if its every edges is properly colored. Consider an  $r$ -coloring of a  $n$ -vertex graph  $H(V, E')$ . For any  $K_k$  in  $H$ ,  $k \in \mathcal{N}$ , a rainbow of size  $x$  exists if there exists a  $K_x$  which is a subgraph of the  $K_k$ ,  $x \leq k$ , and is properly colored. Consider the problem of finding the maximum number of distinct  $K_k$ 's in an  $r$ -coloring such that each  $K_k$  has a rainbow of size  $p$ . It is easy to see that this problem is equivalent to the problem of finding the maximum number of hyperedges in an  $n$ -vertex  $k$ -uniform hypergraph  $G(V, E)$  that can be properly  $(r, p)$  colored by a single  $r$ -coloring: each

$K_k$  is replaced by a  $k$ -uniform hyperedge and a rainbow of size  $p$  denotes  $p$  distinctly colored vertices in the hyperedge.

This problem has been motivated by the separation problems in graphs.

**Definition 1.** Let  $[n]$  denote the set  $1, 2, \dots, n$ . A set  $S \subseteq [n]$  separates  $i$  from  $j$  if  $i \in S$  and  $j \notin S$ . A set  $\mathcal{S}$  of subsets of  $[n]$  is a separator if, for each  $i, j \in [n]$  with  $i \neq j$ , there is a set  $S$  in  $\mathcal{S}$  that separates  $i$  from  $j$ . If, for each  $(i, j) \in [n] * [n]$  with  $i \neq j$ , there is a set  $S \in \mathcal{S}$  that separates  $i$  from  $j$  and a set  $T \in \mathcal{S}$  that separates  $j$  from  $i$ , then  $\mathcal{S}$  is called a complete separator. Moreover, with the additional constraint that the sets  $S$  and  $T$  that separate  $i, j$  are required to be disjoint, then  $\mathcal{S}$  is called a total separator.

We refer the reader to [11, 4, 5, 17, 2, 14, 8] for discussions and results on separating families for graphs. The notion of separation for hyperedges is introduced in [9]. A family  $\mathcal{S} = \{S_1, \dots, S_r\}$  is called a separator for a  $k$ -uniform hypergraph  $G(V, E)$ ,  $S_i \subset V$  for  $1 \leq i \leq r$ , such that every hyperedge  $e \in E$  has a nonempty intersection with at least one  $S_i$  and  $V \setminus S_i$ . We consider the following problem of separation for  $k$ -uniform hypergraphs. Let  $G(V, E)$  be a  $k$ -uniform hypergraph. A set  $\mathcal{S}_1 = \{S_{11}, \dots, S_{1r}\}$   $(r, p)$ -separates a hyperedge  $e \in E$  if (i)  $S_{1j} \subset V$ ,  $S_{1j} \neq \emptyset$ ,  $1 \leq j \leq r$ , (ii)  $\cup_j S_{1j} = V$ , and, (ii)  $e$  has nonempty intersection with at least  $\min\{|e|, p\}$  elements of  $\mathcal{S}_1$ . Observe that the maximum number of hyperedges that can be  $(r, p)$ -separated by a single family  $\mathcal{S}_1$  is  $M(n, k, r, p)$ .

Consider the problem of maximizing profit between a player  $P$  and an adversary  $A$ . Adversary  $A$  provides  $n, k, r$  and  $p$  to the player  $P$ .  $P$  performs some calculation on those parameters and finds out a number  $\#e$ . Now,  $A$  constructs a  $n$ -vertex  $k$ -uniform hypergraph with  $\#e$  hyperedges and colors the vertices with  $r$ -colors. If  $A$  can properly  $(r, p)$  color at least  $\#e$  hyperedges in a hypergraph, then  $A$  wins. If  $A$  cannot properly  $(r, p)$  color at least  $\#e$  hyperedges in a hypergraph, then  $P$  wins. However, the profit of  $P$  is given by  $\binom{n}{k} - \#e$ . So, given a fixed  $n, k, r$  and  $p$ , what value of  $\#e$  should  $P$  use so that he is guaranteed a win and his profit is maximized. Observe that if  $P$  chooses  $\#e$  to be  $M(n, k, r, p) + 1$ , then he is guaranteed a win with maximum profit.

The problem has many applications in resource allocation and scheduling. Consider the problem where there are total  $n$  resources  $\{v_1, \dots, v_n\}$ ,  $m$  processes  $\{e_1, \dots, e_m\}$ . Each process has a distinct wish-list of  $k$  resources. There are  $r$  time slots. A process can execute if it gets at least  $p$  distinct resources in different time slots. The problem is to maximize the number of processes that can be executed within  $r$  time slots. The solution to the above problem is equivalent to the maximum number of hyperedges that can be properly  $(r, p)$  colored by a single  $r$ -coloring in an  $n$ -vertex  $k$ -uniform hypergraph  $G(V, E)$ , where  $V = \{v_1, \dots, v_n\}$ ,  $E = \{e_1, \dots, e_m\}$ . Throughout the paper,  $G$  denotes a  $k$ -uniform hypergraph with vertex set  $V$  and hyperedge set  $E$ , unless otherwise stated.

## 1.1 Motivation

Turán's theorem is a fundamental result in graph theory that gives the maximum number of edges that can be present in a  $K_{r+1}$  free graph. The problem was first stated by Mantel [18, 1] for the special case of triangle free graphs. He proved that the maximum number of edges in an  $n$ -vertex

triangle-free graph is  $\lfloor \frac{n^2}{4} \rfloor$ . Turán [1] posed the same problem for general  $K_{t+1}$ -free graphs and showed that the maximum number of edges in an  $n$ -vertex  $K_{t+1}$ -free graph is  $(1 - \frac{1}{t})\frac{n^2}{2}$ . The graph with  $(1 - \frac{1}{t})\frac{n^2}{2}$  edges is a Turán graph  $T(n, t)$  - a complete  $t$ -partite graph with the size of partite sets differing by at most 1. In the same spirit, Erdős et al.[3] posed the question of maximum number of edges in a graph  $G(V, E)$  that does not contain some arbitrary subgraph  $F$ .

**Definition 2.** Given an  $k$ -uniform hypergraph  $F$  the Turán number  $ex(n, F)$  is the maximum number of hyperedges in an  $n$ -vertex  $k$ -uniform hypergraph not containing a copy of  $F$ . The Turán density  $\pi(F)$  of  $F$  is

$$\pi(F) = \lim_{n \rightarrow \infty} \frac{ex(n, F)}{\binom{n}{k}}.$$

They showed that for a arbitrary graph  $F$  and a fixed  $\varepsilon > 0$ , there exists a  $n_0$  such that for any  $n > n_0$ ,

$$(1 - \frac{1}{\chi(F) - 1} - \varepsilon)\frac{n^2}{2} \leq ex(n, F) \leq (1 - \frac{1}{\chi(F) - 1} + \varepsilon)\frac{n^2}{2},$$

where  $\chi(F)$  denotes the chromatic number of graph  $F$ . For complete graph  $K_{t+1}$ , the chromatic number is  $t + 1$ ; so, the result due to Erdős et al. for  $ex(n, F)$  reduces to an approximate version of Turán's theorem. If  $F$  is bipartite,  $ex(n, H) \leq \varepsilon n^2$ , for  $\varepsilon > 0$ . This also implies that for a graph  $G$ ,  $\pi(F) = (1 - \frac{1}{\chi(F) - 1})$ .

Having solved the problem for  $F = K_t$ , Turán [16] posed the natural generalization of the problem for determining  $ex(n, F)$  where  $F = K_t^k$  is a complete  $k$ -uniform hypergraph on  $t$  vertices. The minimum number of hyperedges in an  $k$ -uniform hypergraph  $G$  on  $n$  vertices such that any subset of  $r$  vertices contains at least one hyperedge of  $G$  is the *Turán number*  $T(n, r, k)$ . Note that  $G$  has this property if and only if the *complementary*  $k$ -uniform hypergraph  $G'$  is  $K_r^k$ -free; thus  $T(n, r, k) + ex(n, K_r^k) = \binom{n}{k}$ . There is extensive study of both  $T(n, r, k)$  and  $ex(n, K_r^k)$  and we refer the reader to two surveys [13, 6] for details. All the above results assumes that the host graph or hypergraph is arbitrary. Mubayi and Talbot [10], and Talbot [15] introduced Turán problems with coloring conditions, which could also be viewed from the perspective of a constrained host graph. They considered a new type of extremal hypergraph problem: given an  $k$ -uniform hypergraph  $F$  and an integer  $r \geq 2$ , determine the maximum number of hyperedges in an  $F$ -free,  $r$ -colorable  $r$ -graph on  $n$  vertices. In similar direction, we pose the following problem: maximize the number of hyperedges in a  $r$ -coloring of a  $n$ -vertex  $k$ -uniform hypergraph  $G$ , such that no hyperedge of  $G$  consists of less than  $p$  colors.

## 1.2 Our Results

In order to estimate  $M(n, k, r, p)$ , we first consider the case when  $r$  divides  $n$  and compute the number of distinct hyperedges that consists of exactly  $p$  distinct colors under any balanced  $r$  coloring of a  $K_n^k$ . Let  $m(n, k, r, p)$  denote the number of distinct hyperedges that consists of exactly  $p$  distinct colors under any balanced  $r$  coloring of a  $K_n^k$ . We prove the following lemma.

**Lemma 1.** For a fixed value of  $n, k, r$  and  $p$ ,  $m(n, k, r, p) = \binom{r}{p} \left( \binom{\frac{n}{r}p}{k} - p \binom{\frac{n}{r}(p-1)}{k} + \binom{p}{2} \binom{\frac{n}{r}(p-2)}{k} \dots (-1)^c \binom{p}{c} \binom{\frac{n}{r}c}{k} \right)$ , where  $c$  is the smallest integer such that  $\frac{n}{r}c \geq k$ .

Observe that summing over all the hyperedges with exactly  $i$  distinct colors,  $1 \leq i \leq p-1$ , we get the number of hyperedges that are colored with at most  $p-1$  colors by any balanced  $r$ -coloring, provided  $r$  divides  $n$ . In Section 3, we show that the number of distinct hyperedges that consists of at least  $p$  distinct colors is maximized when the  $r$ -coloring is balanced. Therefore, we conclude the following theorem.

**Theorem 1.** The maximum number of properly  $(r, p)$  colored hyperedges of a  $K_n^k$  in any  $r$ -coloring (i) is  $M(n, k, r, p) = \binom{n}{k} - \sum_{i=1}^{p-1} m(n, k, r, i)$ , where  $m(n, k, r, i) = \binom{r}{i} \left( \binom{\frac{n}{r}i}{k} - i \binom{\frac{n}{r}(i-1)}{k} + \binom{i}{2} \binom{\frac{n}{r}(i-2)}{k} \dots (-1)^c \binom{i}{c} \binom{\frac{n}{r}c}{k} \right)$ , and,  $c$  is the smallest integer such that  $\frac{n}{r}c \geq k$ , and, (ii) the  $r$ -coloring that maximizes the number of properly colored hyperedges splits the vertex set into equal sized parts, provided  $r$  divides  $n$ .

Furthermore, we generalize the above theorem for arbitrary  $n$  i.e. to cases where  $n$  does not divide  $r$  and derive a upper and lower bound for  $M(n, k, r, p)$  as given by the following theorem.

**Theorem 2.** For a fixed  $n, k, r$  and  $p$ , the maximum number of properly  $(r, p)$  colored  $k$ -uniform hyperedges  $M(n, k, r, p)$  on any  $n$ -vertex hypergraph  $G$  is at most  $\binom{n}{k} - \sum_{i=1}^{p-1} m(n_1, k, r, i)$  and at least  $\binom{n}{k} - \sum_{i=1}^{p-1} m(n_2, k, r, i)$ , where  $n_1 = \lfloor \frac{n}{r} \rfloor \cdot r$ ,  $n_2 = \lceil \frac{n}{r} \rceil \cdot r$ , and  $m(n', k, r, i) = \binom{r}{i} \left( \binom{\frac{n'}{r}i}{k} - i \binom{\frac{n'}{r}(i-1)}{k} + \binom{i}{2} \binom{\frac{n'}{r}(i-2)}{k} \dots (-1)^c \binom{i}{c} \binom{\frac{n'}{r}c}{k} \right)$ , and,  $c$  is the smallest integer such that  $\frac{n'}{r}c \geq k$ . Moreover, the number of properly  $(r, p)$  colored hyperedges is maximized when the  $r$ -coloring is balanced.

### 1.3 Notations

1. For a set  $A$ ,  $\binom{[A]}{r}$  denotes the set of all the distinct  $r$ -element subsets of  $A$ . For instance,  $\binom{[n]}{r}$  denotes the set of all the distinct  $r$ -element subsets of  $\{1, \dots, n\}$ ,  $|\binom{[n]}{r}| = \binom{n}{r}$ .
2. For a set  $S = \{S_1, \dots, S_l\}$ , for any fixed  $l$ ,  $U(S)$  denotes the union of the elements, i.e  $U(S) = S_1 \cup \dots \cup S_l$ .
3. **Lexicographic ordering.** Consider a  $n$ -element set  $V = \{v_1, \dots, v_n\}$  and a set of  $k$ -element subsets  $E = \{e_1, \dots, e_m\}$  of  $V$ , where  $e_i \subset V$ , for  $1 \leq i \leq m$ . For any  $v_q, v_r \in V$ ,  $v_q \prec v_r$  if  $q \leq r$ . Let  $e_i, e_j \in E$ , where  $e_i = \{v_{i1}, \dots, v_{ik}\}$  and  $e_j = \{v_{j1}, \dots, v_{jk}\}$ . Then,  $e_i \prec e_j$  if there exists an index  $l$  such that  $v_{i1} = v_{j1}, \dots, v_{i(l-1)} = v_{j(l-1)}$  and  $v_{il} \prec v_{jl}$ . An ordering  $O$  of subsets of  $E$  is a lexicographic ordering if for every  $e_i, e_j \in O$ ,  $e_i$  precedes  $e_j$  in  $O$  if and only if  $e_i \prec e_j$ .

## 2 Exact Number of properly $(r, p)$ colored hyperedges in a balanced partition

Let  $G(V, E)$  be a  $n$ -vertex  $k$ -uniform hypergraph, where  $V$  denotes the vertex set and  $E$  denotes the set of hyperedges. An  $r$ -coloring  $X$  of vertices in  $V$  partitions the vertex set into  $r$  color

classes  $A = \{A_1, \dots, A_r\}$ , where  $A_j \subseteq V$ ,  $1 \leq j \leq r$  and every vertex  $v \in A_j$  receives the same color under  $X$ . An  $r$ -coloring of vertices is called balanced if every color class is of almost same size, i.e. for all  $A_j \in A$ ,  $|A_j| = \lceil \frac{n}{r} \rceil$  or  $|A_j| = \lfloor \frac{n}{r} \rfloor$ . Let  $p$  be some fixed integer,  $1 < p \leq r$  and  $p \leq k$ . In this section, we study the number of distinct hyperedges that consists of exactly  $p$  distinct colors under any balanced  $r$  coloring of  $G$ . Throughout the section, we assume that  $n$  is divisible by  $r$ , such that for all  $A_j \in A$ ,  $|A_j| = \frac{n}{r}$ .

Consider a balanced  $r$  coloring  $X$  of vertices a  $K_n^k$ . Let  $A = \{A_1, \dots, A_r\}$  denote the corresponding color partition. Let  $m(n, k, r, p)$  denote the number of distinct hyperedges that consists of exactly  $p$  distinct colors under  $X$ . Let  $B$  be the set of all the  $p$ -element subsets  $B_i$  of  $A$ ,  $1 \leq i \leq \binom{r}{p}$  i.e.  $B = \{B_i | B_i \text{ is the } i\text{th } p\text{-element subset of } \binom{[A]}{p}\}$ . Consider the  $i$ th  $p$ -element subset  $B_i \in B$ . Let  $m_i(n, k, r, p)$  denote the number of distinct hyperedges  $e \in E$  that consists of exactly  $p$  distinct colors under  $X$  and  $e \subseteq U(B_i)$ . Due to the balanced nature of the  $r$ -coloring  $X$ , note that  $m_i(n, k, r, p) = m_l(n, k, r, p)$ , for any  $B_i, B_l \in B$ . Observe that

$$m(n, k, r, p) = \sum_{B_i \in B} m_i(n, k, r, p) = \binom{r}{p} m_i(n, k, r, p). \quad (1)$$

So, we focus our attention on computing  $m_i(n, k, r, p)$  for a fixed  $p$ -element subset  $B_i \in B$ . Without loss of generality, we consider  $B_1 = \{A_1, \dots, A_p\}$  as the fixed  $p$ -element subset of  $B$  and compute  $m_1(n, k, r, p)$ .

There are exactly  $p$  subsets of size  $p-1$  of  $B_1$ . Let these sets be  $P_1, \dots, P_p$ , in the lexicographic order. Let  $N(P_j)$  denote the number of hyperedges  $e \in E$  such that  $e \subseteq U(P_j)$ ,  $P_j \in B_1$ , and let  $N(P_j \dots P_l)$  denote the number of hyperedges  $e \in E$  such that  $e \subseteq U(P_j) \cap \dots \cap U(P_l)$  and,  $P_j, \dots, P_l \in B_1$ . Observe that  $N(P_j) = \binom{\frac{n}{r}(p-1)}{k}$ ,  $1 \leq j \leq p$ . So,

$$\sum_{1 \leq j \leq p} N(P_j) = p \binom{\frac{n}{r}(p-1)}{k}. \quad (2)$$

Note that if  $e \subseteq U(P_{j_1})$  and  $e \subseteq U(P_{j_2})$ , then  $e \subseteq U(P_{j_1}) \cap U(P_{j_2})$ . Observe that  $P_{j_1}$  and  $P_{j_2}$  can have at most  $p-2$  parts in common;  $e \subseteq U(P_{j_1} \cap P_{j_2})$  implies that  $e$  lies in a fixed subset of  $p-2$  parts of  $P_{j_1}$ , that is also a subset of  $P_{j_2}$ . So, number of hyperedges  $e$  that lie in a fixed  $p-2$  parts  $P_{j_1} \cap P_{j_2}$  is  $N(P_{j_1} P_{j_2}) = \binom{\frac{n}{r}(p-2)}{k}$ . Since there are exactly  $\binom{p}{2}$  distinct pairs of the form  $\{P_{j_1}, P_{j_2}\}$ , total number of hyperedges  $e$  that are subsets of  $p-2$ -sized subsets of  $B_1$  is

$$\sum_{1 \leq j_1 < j_2 \leq p} N(P_{j_1} P_{j_2}) = \binom{p}{2} \binom{\frac{n}{r}(p-2)}{k}. \quad (3)$$

Let  $c$  be the smallest integer such that  $\frac{n}{r}c \geq k$ . Then,  $\frac{n}{r}(c-1) < k$ . Consider any fixed  $c-1$  parts  $A_{j_1}, \dots, A_{j_{c-1}}$ . Observe that for any hyperedge  $e \in E$ ,  $e \not\subseteq A_{j_1} \cup \dots \cup A_{j_{c-1}}$ . So, we compute all the summations of the form  $\sum_{1 \leq j_1 < j_2 < \dots < j_c \leq p} N(P_{j_1} P_{j_2} \dots P_{j_c})$  till  $\frac{n}{r}c \geq k$ , where

$$\sum_{1 \leq j_1 < j_2 < \dots < j_c \leq p} N(P_{j_1} P_{j_2} \dots P_{j_c}) = \binom{p}{c} \binom{\frac{n}{r}c}{k}. \quad (4)$$

Observe that if a hyperedge  $e$  is a subset of  $B_1$  and is not a subset of any of the  $P_j$ ,  $P_j \in \{P_1, \dots, P_p\}$ , then  $e$  consists of exactly  $p$  colors in the  $r$ -coloring. The total number of hyperedges  $e \subseteq B_1$  is  $N(B_1) = \binom{\frac{n}{r}p}{k}$ . So, by definition,  $N(P'_1, \dots, P'_p)$  denotes all the hyperedges  $e \subseteq B_1$  such that  $e$  consist of exactly  $p$  colors, i.e.  $m_1(n, k, r, p) = N(P'_1, \dots, P'_p)$ . In order to compute  $m_1(n, k, r, p)$ , we use the fundamental result of inclusion exclusion stated below.

**Theorem 3.** [12] *Let  $A$  be any  $n$ -element set, and let  $P_1, \dots, P_m$  denote  $m$  properties of elements of  $A$ . Let  $A_i \subset A$  is the subset of elements of  $A$  with property  $P_i$ . Let  $N(P_i)$  denote the number of elements of  $A$  with property  $P_i$ , i.e.  $N(P_i) = |A_i|$ , for  $1 \leq i \leq m$ . Let  $N(P_i P_j \dots P_l) = |A_i \cap A_j \cap \dots \cap A_l|$ . Let  $N(P'_i)$  denote the number of elements of  $A$  that does not satisfy property  $P_i$  and the number of elements with none of the properties  $P_i, P_j, \dots, P_l$  is denoted by  $N(P'_i P'_j \dots P'_l)$ . Then,*

$$N(P'_1 P'_2 \dots P'_m) = n - \sum_{1 \leq i \leq m} N(P_i) + \sum_{1 \leq i < j \leq m} N(P_i P_j) - \dots + (-1)^m N(P_1 P_2 \dots P_m). \quad (5)$$

So, using principle of inclusion exclusion 3, we have,

$$\begin{aligned} N(P'_1, \dots, P'_p) &= N(B_1) - \sum_{1 \leq j \leq p} N(P_j) + \sum_{1 \leq j_1 < j_2 \leq p} N(P_{j_1} P_{j_2}) - \dots (-1)^c \sum_{1 \leq j_1 < j_2 < \dots < j_c \leq p} N(P_{j_1} P_{j_2} \dots P_{j_c}) \\ &= \binom{\frac{n}{r}p}{k} - p \binom{\frac{n}{r}(p-1)}{k} + \binom{p}{2} \binom{\frac{n}{r}(p-2)}{k} \dots (-1)^c \binom{p}{c} \binom{\frac{n}{r}c}{k}. \end{aligned} \quad (6)$$

Now, using Equation 1, we get,  $m(n, k, r, p) = \binom{r}{p} \left( \binom{\frac{n}{r}p}{k} - p \binom{\frac{n}{r}(p-1)}{k} + \binom{p}{2} \binom{\frac{n}{r}(p-2)}{k} \dots (-1)^c \binom{p}{c} \binom{\frac{n}{r}c}{k} \right)$ , where  $c$  is the smallest integer such that  $\frac{n}{r}c \geq k$ . This concludes the proof of Lemma 1.

Observe that summing over all the hyperedges with exactly  $i$  distinct colors,  $1 \leq i \leq p-1$ , we get the number of hyperedges that are colored with at most  $p-1$  colors by any balanced  $r$ -coloring, provided  $r$  divides  $n$ . Therefore, the exact number of properly  $(r, p)$  colored hyperedges in a balanced partition is

$$M(n, k, r, p) = \binom{n}{k} - \sum_{i=1}^{p-1} m(n, k, r, i). \quad (7)$$

Consider the case when  $r = p = 2$ , i.e., when we are performing a bicoloring on  $n$  vertices and proper coloring of a hyperedge  $e$  denote  $e$  becoming non-monochromatic under the bicoloring. Observe that  $M(n, k, 2, 2) = \binom{n}{k} - m(n, k, 2, 1)$ , and  $m(n, k, 2, 1) = 2 * \binom{\frac{n}{2}}{k}$ . Therefore,  $M(n, k, 2, 2) = \binom{n}{k} - 2 * \binom{\frac{n}{2}}{k}$ , which agrees with the existing results. Note that  $M(n, k, r, p)$  is a non-decreasing function of  $n$ . So,  $M(n-1, k, r, p) \leq M(n, k, r, p) \leq M(n+1, k, r, p)$ .

Let  $x(i, j, n, k, r) = \binom{r}{i} \binom{\frac{n}{r}i}{k} - \frac{r-j}{i-j+1} x(i, j-1, n, k, r)$ .  $x(i, j, n, k, r)$  denotes the number of hyperedges that are colored with less than or equal to  $j$  colors by an  $r$ -coloring, when counted with respect to color classes of size  $i$ ,  $i \geq j$ . Here, the term  $\binom{r}{i} \binom{\frac{n}{r}i}{k}$  accounts for every hyperedge  $e \in E$ , that is a subset of some fixed  $i$  color parts of the  $r$ -coloring. Any  $(j-1)$ -sized color parts are repeated  $r-j+1$  times when counted over all  $j$ -sized color classes; however, we need to count it exactly once. Each hyperedge inside some fixed  $i$ -sized set is counted  $i-j+1$  times over all the  $j-1$  sized sets. So,  $\binom{r}{i} \binom{\frac{n}{r}i}{k} - \frac{r-j+1}{i-j+1} x(i, j-1, n, k, r) + \frac{1}{i-j+1} x(i, j-1, n, k, r)$  counts the number of hyperedges that are colored with less than or equal to  $j$  colors by an  $r$ -coloring,

when counted with respect to color classes of size  $i$ ,  $i \geq j$ .  $\frac{1}{i-j+1}x(i, j-1, n, k, r)$  term is added in order to include the hyperedges colored with less than or equal to  $j-1$  colors. Observe that  $x(p-1, p-1, n, k, r)$  denotes the number of hyperedges colored with less than or equal to  $p-1$  colors by a balanced  $r$ -coloring. Therefore,

$$M(n, k, r, p) = \binom{n}{k} - x(p-1, p-1, n, k, r). \quad (8)$$

### 3 Maximizing the number of properly $(r, p)$ colored hyperedges

In this section, we show that the number of properly  $(r, p)$  colored hyperedges is maximized when the  $r$ -coloring is balanced. We show that the number of hyperedges colored with less than or equal to  $p-1$  colors is minimized for a balanced  $r$ -coloring, thereby proving the above claim.

Consider an  $r$ -coloring  $X$  of vertices a  $K_n^k$ . Let  $A = \{A_1, \dots, A_r\}$  denote the corresponding color partition and let  $|A_i| = n_i$ , for  $1 \leq i \leq r$ . Let  $m_X(n, k, r, p)$  denote the number of distinct hyperedges that consists of at most  $p$  distinct colors under  $X$ . Let  $n_1 \geq n_2 + 2$ . Then we have the following lemma.

**Lemma 2.** *The number of hyperedges colored with at most  $p$  colors is reduced by moving a vertex  $v \in A_1$  from  $A_1$  to  $A_2$ , i.e. switching the color of  $v$  from 1 to 2 produces an  $r$ -coloring  $X'$  such that  $m_{X'}(n, k, r, p) < m_X(n, k, r, p)$ .*

**Proof** In order to prove that  $m_{X'}(n, k, r, p) < m_X(n, k, r, p)$ , we analyze: (i) the *gain*  $g$ : the number of hyperedges  $e \in E$  such that  $e$  is colored with greater than  $p$  colors under  $X$  and  $e$  receives at most  $p$  colors under  $X'$ , and, (ii) the *loss*  $l$ : the number of hyperedges  $e \in E$  such that  $e$  is colored with at most  $p$  colors under  $X$  and  $e$  receives at least  $p+1$  colors under  $X'$ . Note that a hyperedge  $e \in E$  contributes to  $g$  or  $l$  if and only if  $v \in e$ . Since  $m_{X'}(n, k, r, p) = m_X(n, k, r, p) + g - l$ , in order to prove Lemma 2, we need to show that  $l > g$ .

Let  $y(n, k, r, p)$  denote the minimum number of  $k$ -uniform hyperedges on  $n$  labeled vertices that are colored with exactly  $p$  colors by any  $r$  coloring. Observe that a hyperedge  $e \in E$  contributes to  $g$  if and only if it consists of exactly  $p+1$  colors in  $X$ ,  $v \in e$  and includes no other vertex from  $A_1$ , i.e.,  $e \cap A_1 = v$ , and includes at least one vertex from  $A_2$ , i.e.,  $e \cap A_2 \geq 1$ . So, gain due to switching  $v$  from  $A_1$  to  $A_2$  is

$$g = \sum_{i=1}^c \binom{n_2}{i} y(n - n_1 - n_2, k - i - 1, r - 2, p - 1), \quad (9)$$

where  $c$  be the smallest integer such that  $\frac{n}{r}c \geq k$ . In each of the  $c$  terms in the summation,  $\binom{n_2}{i}$  denotes the number of ways to choose exactly  $i$  vertices from  $A_2$  (of color 2),  $y(n - n_1 - n_2, k - i - 1, r - 2, p - 1)$  denotes the minimum number of hyperedges that can be formed consisting of exactly  $k - (i + 1)$  vertices from  $A \setminus (A_1 \cup A_2)$  and exactly  $p - 1$  distinct colors. The  $k - (i + 1)$  vertices from  $A \setminus (A_1 \cup A_2)$  with  $p - 1$  distinct colors combined with  $i$  vertices from  $A_2$  and  $v$  from  $A_1$  forms the hyperedges  $e$  consisting of exactly  $p + 1$  colors under coloring  $X$  including  $v$ ,  $e \cap A_1 = v$ , and  $|e \cap A_2| = i$ .

Similarly, a hyperedge  $e \in E$  contributes to  $l$  if and only if it consists of exactly  $p$  colors in  $X$ , includes no other vertex from  $A_2$ , i.e.,  $e \cap A_2 = \emptyset$ , and  $v \in e$  and includes at least one vertex other than  $v$  from  $A_1$ , i.e.,  $|e \cap A_1| \geq 2$ . So, loss due to switching  $v$  from  $A_1$  to  $A_2$  is

$$l = \sum_{i=1}^c \binom{n_1-1}{i} y(n-n_1-n_2, k-i-1, r-2, p-1). \quad (10)$$

Since  $n_1 \geq n_2 + 2$ ,  $n_1 - 1 > n_2$ . So, comparing  $l$  and  $g$  term-wise, we get  $l > g$  as desired.  $\square$

Lemma 2 implies that the number of hyperedges colored with less than  $p$  colors can be minimized until the color partition  $\{A_1, \dots, A_r\}$  is balanced, i.e. for every  $i$ ,  $1 \leq i \leq r$ ,  $\lfloor \frac{n}{r} \rfloor \leq |A_i| \leq \lceil \frac{n}{r} \rceil$ . Therefore, the number of properly  $(r, p)$  colored hyperedges is maximized when the  $r$ -coloring is balanced. So, using Equation 7, Theorem 1 follows.

Observe that even if  $r$  does not divide  $n$ , the  $r$ -coloring that maximizes the number of properly colored hyperedges splits the vertex set into almost equal sized parts (from Lemma 2) of either  $\lfloor \frac{n}{r} \rfloor$  or  $\lceil \frac{n}{r} \rceil$  size. Therefore, we can get an upper bound on  $M(n, k, r, p)$  by computing the minimum number of hyperedges including vertices of at most  $p-1$  distinct colors with  $\lfloor \frac{n}{r} \rfloor \cdot r$  vertices and subtracting from  $\binom{n}{k}$ . Furthermore, we can get a lower bound on  $M(n, k, r, p)$  by computing the minimum number of hyperedges including vertices of at most  $p-1$  distinct colors with  $\lceil \frac{n}{r} \rceil \cdot r$  vertices and subtracting from  $\binom{n}{k}$ . This observation combined with Theorem 1 proves Theorem 2.

For the special case when  $r = p = k$ , we can compute  $M(n, k, r, p)$  much easily. Observe that any hyperedge must contain one vertex each from each of the color classes  $\{A_1, \dots, A_r\}$  in order to be properly  $(r, p)$  colored. So, the number of properly colored hyperedges under any  $r$ -coloring is  $|A_1||A_2|\dots|A_r|$ . Using the second part of Theorem 2,  $M(n, k, r, p) = |A_1||A_2|\dots|A_r|$ , where  $\{A_1, \dots, A_r\}$  is a balanced partition. So, we have the following corollary.

**Corollary 1.** *The number of properly  $(r, p)$  colored hyperedges of a  $K_n^k$  in any  $r$ -coloring is  $|A_1||A_2|\dots|A_r|$  when  $r = p = k$ . Moreover, the  $r$ -coloring that maximizes the number of properly colored hyperedges splits the vertex set into almost equal sized parts.*

## References

- [1] Martin Aigner and Günter M. Ziegler. *Proofs from THE BOOK*, pages 235–240. Springer-Verlag, New York, 1998.
- [2] T.J. Dickson. On a problem concerning separating systems of a finite set. *J. Combin. Theory*, 7:191–196, 1969.
- [3] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.
- [4] G. Katona. On separating systems of a finite set. *J. Combin. Theory I*, pages 174–194, 1966.
- [5] G. Katona. Chapter 23 - combinatorial search problems. In Jagdish N. Srivastava, editor, *A Survey of Combinatorial Theory*, pages 285 – 308. North-Holland, 1973.



- [6] Peter Keevash. Hypergraph turan problems. *Surveys in Combinatorics*, pages 83–140, 2011.
- [7] László Lovász. Coverings and colorings of hypergraphs. In *Proc. 4th Southeastern Conference on Combinatorics, Graph Theory, and Computing*, Utilitas Mathematica Publishing, Winnipeg, pages 3–12, 1973.
- [8] CAI Mao-cheng. Solutions to Edmond’s and Katona’s problems on families of separating subsets. *Discrete Mathematics*, 47:13–21, 1983.
- [9] Tapas Mishra and Sudebkumar Prasant Pal. Bicoloring covers for graphs and hypergraphs. *arXiv preprint arXiv:1501.00343*, 2015.
- [10] D. Mubayi and J. Talbot. Extremal problems for  $t$ -partite and  $t$ -colorable hypergraphs. *Electronic J. Combin.*, 15, 2008.
- [11] A. Rényi. On random generating elements of a finite boolean algebra. *Acta Sci. Math. (Szeged)*, 22(1-2):75–81, 1961.
- [12] Kenneth H. Rosen. *Discrete Mathematics and Its Applications*. McGraw-Hill Higher Education, 5th edition, 2002.
- [13] Alexander Sidorenko. What we know and what we do not know about turán numbers. *Graphs and Combinatorics*, 11(2):179–199, 1995.
- [14] J. Spencer. Minimal completely separating systems. *J. Combin. Theory*, 8:446–447, 1970.
- [15] John Talbot. Chromatic turán problems and a new upper bound for the turán density of. *European Journal of Combinatorics*, 28(8):2125 – 2142, 2007. EuroComb ’05 - Combinatorics, Graph Theory and Applications EuroComb ’05 - Combinatorics, Graph Theory and Applications.
- [16] P. Turán. Research problems. *MTA Mat.Kutató Int.Közl.*, 6:417–423, 1961.
- [17] I. Wegener. On separating systems whose elements are sets of at most  $k$  elements. *Discrete Math.*, 28:219–222, 1979.
- [18] Douglas Brent West. *Introduction to Graph Theory*, pages 41–42. Prentice Hall, 2nd edition, 2001.